Functional voice outcomes after thyroidectomy: An assessment of the Dsyphonia Severity Index (DSI) after thyroidectomy

Leonard R. Henry, MD,[a,b,c] **Leah B. Helou, MA, CCC-SLP,**[d] **Nancy Pearl Solomon, PhD, CCC-SLP,**[e] **Robin S. Howard, MA,**[f] **Joyce Gurevich-Uvena, MA, CCC-SLP,**[e] **George Coppit, MD,**[g] *and* **Alexander Stojadinovic, MD,**[b,c,d] *Bethesda, MD, and Washington, DC*

Background. *The Dysphonia Severity Index (DSI) is an objective multiparametric acoustic calculation of vocal function; however, its changes after thyroidectomy have not yet been described.*
Methods. *Patient-reported symptoms, as well as auditory perceptual, acoustic, and videolaryngostroboscopic (VLS) data, were collected prospectively before and after thyroidectomy. Voice outcomes (normal versus negative voice outcome [NVO]) at 6 months after thyroidectomy were based on a combination of voice symptoms and objective findings. The DSI was assessed over the perioperative course, and differences were determined with Wilcoxon signed rank tests. The DSI was compared between study groups (normal versus NVO) using t tests, analyses of variance (ANOVAs), or rank sum tests as appropriate. The predictive value of DSI for long-term voice dysfunction was assessed by an area under the receiver operating characteristics curve analysis. Correlations between DSI and Consensus Auditory Perceptual Ratings of Voice (CAPE-V) and the patient reported Voice Handicap Index (VHI) were determined with Pearson's correlation coefficients.*
Results. *In all, 62 patients were evaluated before, 1–4 weeks after, and 6 months after thyroidectomy. Eight (13%) patients were diagnosed with NVO at 6 months. The DSI was different postoperatively between NVO and normal voice (*P *= .005, repeated measures [RM]-ANOVA), with the NVO group demonstrating a lesser DSI value and greater change from pre-operative assessment at the first postoperative visit when compared with the normal group (*P *< .006 each). The DSI differed significantly for pre-operative and 6-month assessments according to sex, smoking status, and age. Short-term postoperative DSI (area under the curve [AUC] = 0.795) and DSI change from baseline to 1–4 weeks (AUC = 0.835) were highly predictive of 6-month NVO. DSI measurements over the post-thyroidectomy course were correlated poorly to moderately (maximum r = –0.62) with CAPE-V and VHI assessments for the same time points.*
Conclusion. *The DSI is decreased in the early post-thyroidectomy period, mostly in persons who were ultimately found to have a long-term NVO. Early postoperative DSI and change of DSI from baseline at 1–4 weeks postoperation predict long-term post-thyroidectomy voice dysfunction. The modest correlations between the DSI and other vocal assessments point to the utility of DSI as an independent predictor of voice dysfunction after thyroidectomy, which can select patients who may benefit from voice therapy. (Surgery 2010;■:■-■.)*

From the Department of Surgery, National Naval Medical Center,[a] *Bethesda; the Uniformed Services University of the Health Sciences,*[b] *Bethesda, MD; United States Military Cancer Institute,*[c] *Washington, DC; Department of Surgery,*[d] *Army Audiology and Speech Center,*[e] *Washington, DC; the Department of Clinical Investigation, Division of Biostatistics,*[f] *and the Department of Surgery, Division of Otolaryngology–Head and Neck Surgery,*[g] *Walter Reed Army Medical Center, Washington, DC*

L.B.H. is now with the Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA.

Presented in part at the American Association of Endocrine Surgeons Annual Meeting.

Accepted for publication November 24, 2009.

Reprint requests: Leonard R. Henry, MD, Department of Surgery, Uniformed Services University of the Health Sciences, National Naval Medical Center, 8901 Wisconsin Avenue, Bethesda, MD 20889. E-mail: leonard.henry@med.navy.mil.

0039-6060/$ - see front matter

doi:10.1016/j.surg.2009.11.017

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Functional voice outcomes after thyroidectomy: An assessment of the Dsyphonia Severity Index (DSI) after thyroidectomy**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Uniformed Services University of the Health Sciences,National Naval Medical Center,8901 Wisconsin Avenue,Bethesda,MD,20889**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Altered voice is a common problem after thyroid surgery. Previous studies demonstrate that 25% to almost 90% of patients report abnormal voice within the first few weeks after operation, and 11–15% of patients report persistent voice problems 3–6 months after thyroidectomy.[1-8] Certainly, major laryngeal nerve injuries account for a large portion of this incidence and are well-established and feared complications of thyroidectomy,[9-12] but many patients suffer long-term negative voice outcomes (NVOs) and have no evidence of laryngeal nerve injury.[13] Various non-neurogenic mechanisms have been postulated to account for post-thyroidectomy voice changes, including the effects of endotracheal intubation[14] and other alterations in normal anatomy and mechanical factors.[1,3,15]

The ability to identify early changes in voice indicative of durable dysfunction (both neurogenic and non-neurogenic) would facilitate early referral for comprehensive voice evaluation aimed to improve quality of life, prevent secondary injuries, and identify patients who might benefit from vocal fold augmentation. Predictive abilities of all postoperative surveillance strategies (the optimal 1 yet to be defined) remain under the influence of a myriad of factors, which include the reliance on the patient's report of symptoms, the surgeon's awareness of dysphonia, and his or her auditory perception and expertise in interpreting laryngoscopic findings. In contrast, acoustic measurements are objective and relatively easy to obtain and interpret.

Wuyts et al[16] developed a multivariate model of acoustic measures obtained from patients with and without vocal pathology, in which parameters were weighted to create an index of vocal function. This objective measure, which was termed the Dysphonia Severity Index (DSI), incorporated measurements of maximum phonation time (MPT, in seconds), highest vocal frequency (F0-high, Hertz), lowest vocal intensity (I-low, decibels) and percent jitter (a measure of vocal frequency variability). The DSI is scored from –5 to +5, where +5 indicates a normal, healthy voice, and –5 represents a severely pathologic voice. DSI measures have been shown to be highly related to clinician-perceived and patient-reported severity of voice handicap.[16,17] Similarly, the DSI has been used clinically as a tool to evaluate the effectiveness of voice therapy and training, as well as to compare voice quality among groups with laryngeal pathology.[17-20] The DSI decreases with age,[21] which reflects the natural deterioration of vocal function with increasing age. The DSI can identify patients with dysphonia caused by both organic and functional voice disorders and to discriminate voice pathology resulting from lesions or vocal fold paralysis or paresis.[17] The DSI has high interobserver reliability,[22] and test–retest variability results confirm that changes in DSI >2.49 within a given patient are clinically important.[22]

Given the assumption that the DSI is an objective, reliable, and reproducible assessment of voice quality across clinicians and centers, the application of DSI to evaluate and quantify voice disturbances after thyroidectomy is appealing conceptually. The goal of this study was to assess vocal function using the DSI and to determine the range of quantifiable changes that can be expected in patients undergoing thyroidectomy. In addition, we aimed to determine whether the DSI has clinical utility for identifying patients likely to have long-term vocal problems after thyroidectomy. Finally, we correlated the DSI with expert perceptual assessments of dysphonia (Consensus Auditory Perceptual Evaluation of Voice [CAPE-V])[23] and patient reported assessment of perceived voice handicap (Voice Handicap Index [VHI]).[24]

METHODS

Participants. Patients with benign thyroid pathology or primary nonanaplastic thyroid carcinoma, who were scheduled to undergo thyroid resection under general or regional anesthesia, were included in this study approved by the Institutional Review Board of the Walter Reed Army Medical Center. Those patients who had undergone thyroidectomy or other neck operation previously, as well as patients with a documented voice disorder, were excluded.

Recorded data. At each of 4 time points (prethyroidectomy, 1–4 weeks (mean, 14 days), 3 months, and 6 months post-thyroidectomy), participants underwent multiparametric voice testing. Each testing session included high-quality voice recording, collection of acoustic measures, voice expert perceptual analysis of voice quality employing the CAPE-V to be used for acoustic and auditory-perceptual analysis, assessment of laryngeal structure and function via indirect videolaryngostroboscopy (VLS), and collection of patient report of perceived voice handicap using the VHI.

Acoustic recordings were collected by an experienced speech-language pathologist (SLP) using a commercially available acoustic analysis system (Computerized Speech Lab 4500; KayPENTAX, Lincoln Park, NJ). Jitter obtained from a 2-s steady-state portion of a 3–4 s sustained /a/ produced with habitual voice quality, and the longest of 3 maximum phonation trials were included in

the DSI equation. In addition, the highest frequency (Hz) and lowest intensity (dB) were elicited by the clinician across several trials to ensure that the production was representative of the subject's ability at each time point. Measures for each parameter were entered into a database, and the DSI was calculated as described previously by Wuyts.[16]

An automated version of the CAPE-V was used to present all voice recordings in a randomized, blinded fashion.[25] Listeners were exposed to sustained vowels (/i/ and /a/) and the 6 sentences from the CAPE-V in any order and as many times as desired; the ratings were made by clicking on a defined linear region for each CAPE-V parameter. The voice samples were grouped by sex and presented to the listener with a sex-matched auditory anchor that was used at each listener's discretion. Three SLPs with experience in voice disorders rated all voice samples, and the median of the ratings from all 3 SLPs for each parameter of the CAPE-V were recorded as the auditory perceptual measure of voice quality.

At each of the study time points, the patient's perceived impact of voice handicap was reported by completion of the VHI. The VHI is a 30-question validated assessment of vocal handicap, which is stratified into the domains of physical, functional, and emotional impact of voice deficit, as determined by the individual patient.

Imaging of the larynx during phonation and nonphonatory tasks was performed with a fiberoptic flexible laryngoscope passed transnasally, using halogen and stroboscopic light sources. A standardized vocal task protocol during VLS was employed. Patients who demonstrated notable laryngeal abnormalities were diagnosed by a consensus committee of SLPs, surgical oncologists, and/or otolaryngologists, all of whom had experience in voice disorders.

Criteria for NVO. The combined clinical expertise and experience of our research team (3 SLPs, 2 surgical oncologists, and 2 otolaryngologists—all with substantial clinical and research voice experience), the review of evidence-based research, and the theoretic understanding of voice disorders were incorporated to define voice-specific outcomes. NVO was classified as objective or subjective. Those patients included in the objective NVO category demonstrated structural or physiologic laryngeal abnormalities evident on VLS, combined with an increase in VHI score ≥20 points from preoperative baseline and/or an increase in overall severity on the CAPE-V ≥11 points from baseline. Subjective NVO was assigned in the absence of VLS anatomic or physiologic derangement when 2 or more of the following criteria were met: VHI increase ≥20 points from preoperative baseline, patient report of atypical voice compared with pre-operative status, and CAPE-V overall severity increase ≥11 points from preoperative values. All patients not classified as having a subjective or objective NVO were categorized as normal.

Table I. Study group characteristics

	Number (%)
Age, mean years ± SD	51 ± 13
Volume thyroid resected, median cm^3, (range)	63 cm^3 (5–2,872)
Maximum tumor size, median cm (range)	1.3 (0.1–8.0)
Sex	
Male	28 (45%)
Female	34 (55%)
Ethnicity	
Caucasian	41 (66%)
African American	15 (24%)
Other	6 (10%)
Smoker	
Never	41 (66%)
Former	18 (29%)
Current	3 (5%)
Voice professional	
Yes	15 (24%)
No	47 (76%)
Procedure	
Total	37 (60%)
Lobe	25 (40%)
Histology: malignant	
Yes	25 (40%)
No	37 (60%)

SD, Standard deviation.

Statistical methods. Comparisons of DSI changes within groups over time were assessed with Wilcoxon signed rank tests. Comparisons of DSI and cohort characteristics, as well as comparisons of raw DSI scores and DSI changes from preoperative baseline at the specified time points in patients with and without NVO, were made with *t* tests, ANOVAs, or Wilcoxon rank sum tests as appropriate. The estimated areas under the curve (AUCs) of the receiver operating characteristics (ROC) curves were determined for the DSI at 1–4 weeks postoperation and for the DSI change from preoperative baseline at 1–4 weeks in predicting long-term NVO at 6 months after thyroidectomy. The correlations of DSI with CAPE-V and VHI measurements at the specified time points were assessed with Pearson coefficients. Statistical tests were conducted with SPSS software version

Table II. Study group characteristics and DSI at baseline and 6 months after thyroidectomy

	Pre-operative DSI	P *value*	*6-month DSI*	P *value*	*Change*	P *value*
Sex						
Male	–0.12 (± 3.24)	<.005*	–0.24 (± 2.55)	<.005*	–0.11(± 2.40)	NS
Female	2.62 (± 2.34)		2.31 (± 2.24)		–0.30(± 1.88)	
Ethnicity						
Caucasian	1.19 (± 3.17)	NS	1.14 (± 2.84)	NS	–0.04 (± 1.86)	NS
African American	1.38 (± 3.36)		0.86 (± 2.33)		–0.52 (± 2.79)	
Other	2.73 (± 1.21)		2.06 (± 2.67)		–0.67 (± 2.04)	
Smoker						
Current	–1.22 (± 2.18)	.03†	–0.16 (± 2.05)	.003†	1.06 (± 0.16)	NS
Former	0.22 (± 3.84)		–0.45 (± 2.25)		–0.67 (± 2.80)	
Never	2.08 (± 2.52)		1.97 (± 2.59)		–0.11 (± 1.80)	
Voice professional						
Yes	1.21 (± 3.39)	NS	1.78 (± 3.13)	NS	0.56 (± 2.04)	.10
No	1.43 (± 3.01)		0.97 (± 2.54)		–0.47 (± 2.10)	
Procedure						
Total	1.37 (± 2.98)	NS	1.41 (± 2.45)	NS	0.04 (± 2.06)	NS
Lobe	1.41 (± 3.29)		0.81 (± 3.02)		–0.60 (± 2.18)	
Histology: malignant						
Yes	1.31 (± 2.95)	NS	0.98 (± 2.03)	NS	–0.33 (± 2.37)	NS
No	1.43 (± 3.21)		1.29 (± 3.08)		–0.14 (± 1.95)	
Histology: benign inflammatory						
Yes	1.62 (± 3.06)	NS	1.47 (± 2.90)	.11	–0.15 (± 1.93)	NS
No	0.80 (± 3.14)		0.42 (± 1.96)		–0.38 (± 2.56)	
Histology: benign noninflammatory						
Yes	1.86 (± 2.55)	NS	2.10 (± 2.24)	.12	0.25 (± 1.69)	NS
No	1.23 (± 3.24)		0.86 (± 2.77)		–0.37 (± 2.23)	

*Two-sample *t* test.
†ANOVA.
NS, Nonsignificant, $P > .2$.

Table III. Correlations between study group characteristics and DSI at baseline and 6 months after thyroidectomy

	Pre-operative DSI	P *value*	*6-month DSI*	P *value*	*Change*	P *value*
Age	–0.49	<.005*	–0.48	<.005*	0.10	.44
Volume thyroid resected (cm^3)	–0.16	.22	–0.06	.65	0.15	.24
Maximum tumor size (cm)	–0.04	.83	–0.20	.34	–0.12	.59

*Pearson's correlation coefficient.

16 (SPSS Inc., Chicago, IL). For all tests, $P < .05$ was considered significant.

RESULTS

In all, 62 patients who underwent thyroidectomy were followed with comprehensive voice assessment before, 1–4 weeks post-thyroidectomy, and 6 months post-thyroidectomy. Eight of 62 (13%) patients were classified at 6 months as having NVO; of these patients, 6 had subjective NVO and 2 had objective NVO. Both patients in the objective NVO group demonstrated persistent signs of recurrent laryngeal nerve paralysis. Subjective and objective NVO patients were grouped into 1 NVO cohort for statistical analysis. The clinical characteristics of the study population are summarized in Table I. Changes in DSI at 6 months post-thyroidectomy relative to the pre-operative baseline were compared across the clinical characteristics of the study population; significant differences according to their sex, smoking history, and age were revealed (Tables II and III). Women, patients who had never smoked, and younger patients had a higher (better) DSI at the pre-operative and 6 month postoperative time points, but the degree of change from baseline was not different between these groups.

The DSI was less in the entire cohort at the early postoperative visit ($P = .008$), after which changes from baseline were not significant. When separated

Table IV. Values of DSI* over the course of thyroidectomy (pre-operatively to 6 months postoperatively)

	DSI			
	Pre-operatively	*1–4 weeks*	*3 months*	*6 months*
Normal	2.05 (–6.40–6.83)	0.88 (–8.77–7.90)	0.94 (–4.54–5.38)	1.17 (–6.21–6.73)
NVO	1.45 (–1.94–3.93)	–4.30 (–19.10–1.90)	–0.89 (–3.43–2.25)	0.04 (–1.65–2.24)
		Change in DSI		
		1–4 weeks†	*3 months*	*6 months†*
Normal		–0.64 (–14.72–5.05)	–0.70 (–4.63–4.23)	–0.21 (–6.36–4.81)
NVO		–5.75 (–18.87–1.65)	–2.37 (–4.83–2.00)	–1.19 (–4.54–1.19)
P value	.75	.006*	.10	.15

*Median (range) per group.
†Wilcoxon rank sum test.

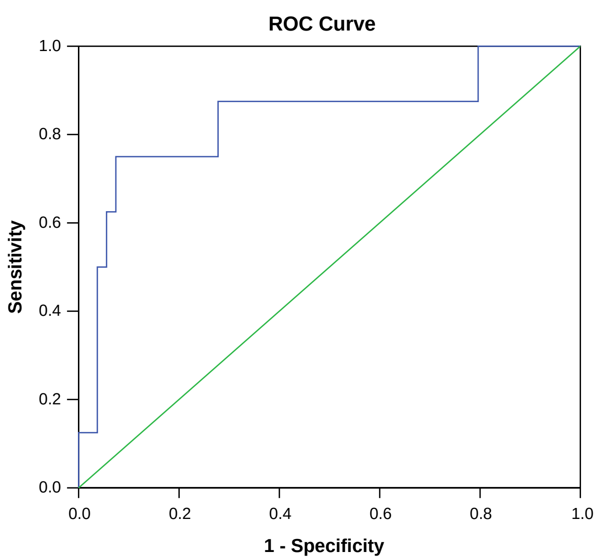

Fig 1. ROC curve (AUC = 0.835) for DSI change at 1 week from baseline in predicting an NVO at 6 months (P = .002).

by group, however, only the NVO group changed significantly from baseline. The DSI over the post-operative course in patients with and without NVO were different from each other (P < .005) with regard to the absolute DSI score at 1 week post-thyroidectomy (Table IV; P = .006), and with regard to the degree of change in DSI at 1 week from baseline (Table IV; P = .001).

Both the absolute DSI score at 1–4 weeks after thyroidectomy and the DSI change from pre-operative baseline at the first postoperative visit predicted NVO at 6 months with area under the ROC curve of 0.795 (P = .008) and 0.835 (P = .002), respectively. Figure 1 shows the ROCs for DSI change from baseline at 1–4 weeks after thyroidectomy. The sensitivity, specificity, and positive and negative predictive values of various DSI changes from pre-operative baseline thresholds (–1.5 to –3.5) to 1–4 weeks after thyroidectomy are shown in Fig 2. Lesser changes from baseline at 1–4 weeks post-thyroidectomy are more sensitive for identifying long-term NVO, whereas greater changes from the baseline are more specific.

Absolute DSI measurements and changes in DSI from pre-operative baseline at first postoperative visit correlated with VHI and CAPE-V measurements of overall severity at the pre-operative and postoperative time points (Table V). Several statistically significant correlations were identified, with the strongest of these being an inverse correlation between the change in total VHI score at 1–4 weeks from pre-operative baseline to the calculated change in DSI at 1–4 weeks (r = –0.62, P < .005).

DISCUSSION

Early recognition and intervention for durable voice problems after thyroid surgery can lead to efforts aimed at improving quality of life, preventing maladaptive behaviors that might produce secondary injury, and identifying appropriate patients for consideration for vocal fold augmentation to improve voice and mitigate aspiration risk. The ideal surveillance strategy to identify patients at risk of long-term voice problems after thyroidectomy, however, remains elusive. The DSI represents a relatively easily obtained and serviceable measurement of voice toward that end. To our knowledge, this article is the first to evaluate the DSI in the context of post-thyroidectomy functional voice assessment. In this study, we determined the DSI in a cohort of patients who underwent thyroidectomy and explored the utility of early postoperative DSI changes (determined at 1–4 weeks, relative to pre-operative baseline) in predicting long-term NVOs (at the sixth postoperative month). In addition, we evaluated whether the raw (ie,

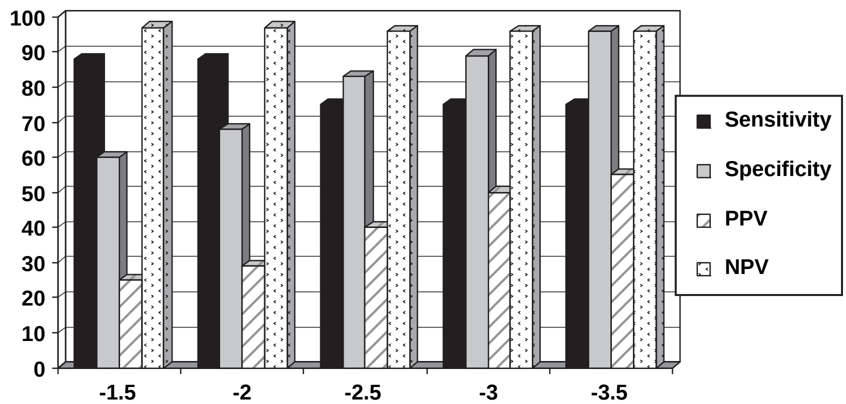

Fig 2. Sensitivity, specificity, and positive and negative predictive values (y-axis) of DSI change at 1–4 weeks from baseline (x-axis) in predicting an NVO at 6 months post-thyroidectomy.

Table V. Pearson's correlation coefficients pre-operatively at 1 week, 3 months, and 6 months postoperatively (*P* value) between absolute CAPE-V/Total VHI and DSI

	Pre DSI	*Post1 DSI*	*Post2 DSI*	*Post3 DSI*
CAPE-V severity	–0.295 (0.02)	–0.528 (<.005)	–0.255 (.055)	–0.322 (.011)
Total VHI	–0.100 (0.44)	–0.620 (<.005)	–0.235 (.076)	–0.198 (.12)
		Post1 DSI change	*Post2 DSI change*	*Post3 DSI change*
CAPE-V severity change		–.580 (<0.005)	–.329 (.012)	–.180 (.17)
Total VHI change		–.601 (<0.005)	–.272 (.039)	–.211 (.099)

Voice outcomes assessment tools and NVO.

absolute) DSI scores or changes from preoperative baseline DSI scores determined during the postoperative follow-up period correlated with other measures related to vocal outcome currently in clinical use (ie, the patient-reported VHI and perceptual ratings of voice quality via the CAPE-V).

Age, sex, and smoking status were identified as clinically relevant determinants of DSI in this patient population. At the first postoperative visit, the absolute DSI and the change from preoperative baseline were both significantly different in patients with normal voice versus those experiencing persistent NVO 6 months after thyroidectomy. Furthermore, DSI change from baseline to the first postoperative visit after thyroidectomy was predictive of which patients would experience durable NVO at 6 months post-thyroidectomy (AUC = 0.835; *P* = .002).

Currently, there is no consensus regarding what constitutes an NVO after thyroidectomy. Although laryngeal nerve injuries are considered widely to be secondary surgical events, the non-neurologic causes of impaired voice after thyroid surgery are often unrecognized. A global definition of subjective and objective NVOs has yet to be established. Thus, patient complaints that cannot be linked to physiologic impairment are often not reported in the literature or are variably classified. In this study, we have also introduced a multifactorial definition for subjective and objective NVO that is based on years of collective clinical and research experience.

Age is known to affect the DSI,[21] and we observed the same phenomenon in this study. In addition, patients who had never smoked showed higher (better) DSI values at the pre-operative session, as well as significantly higher DSI calculations at the 6-month postoperative time point. These results are intuitive, because smoking is known to decrease maximum phonation time[26] and is an established contributor to dysphonia and other laryngeal pathology.[27] The DSI was shown originally to be independent of sex, supposedly because the greater maximum pitch of the female voice is offset by the greater maximum phonation time achieved by men.[16] Our findings are at odds with this claim, as the DSI was significantly greater in women than men at the preoperative and 6-month assessments, although the degree of change in DSI from baseline to any postoperative time point was not different across sex. One explanation for this

Table VI. Pearson's correlation coefficients of changes from baseline in CAPE-V/Total VHI and DSI at each postoperative time point

Abbreviation	*Full term*	*Description*	*Benefits*	*Limitations*
DSI	Dysphonia Severity Index	An acoustic calculation of voice function usually measured from –5.0 to 5.0. Changes from baseline of 2.5 or greater are believed to be clinically significant	Objective, easily obtained, reproducible and portable; indicated to gauge pathology or document response to therapy	Relies on maximal-performance tasks that depend on subject's motivation and ability Requires high-quality acoustic recording equipment and a quiet recording environment
VLS	Videolaryngoscopy or Videolaryngostroboscopy	Recorded examination of larynx during functional activities using both halogen and stroboscopic light sources	Recorded for repeated review. Affords detailed examination of laryngeal structure and function to permit anatomic and functional diagnoses when present	Limited by subjective interpretations of sometimes subtle findings; stroboscopic light source can induce optical illusions; light imbalances can affect ratings
CAPE-V	Consensus Auditory Perceptual Evaluation of Voice	Current rating of voice quality as determined by skilled listener. Ratings of overall severity, roughness, breathiness, strain, pitch, and loudness on a modified visual analog scales	Popular assessment for diagnosis of "dysphonia"; developed by a consensus panel of experts.	Best reserved for employment by voice professionals. Less reproducible and lower inter-rater reliability among non-voice experts. Few studies have tested its reliability to date
VCH	Voice Case History	Collection of patient-reported specific voice symptoms: voice typical, difficulty singing, difficulty projecting, hoarseness, vocal fatigue, and pain	Easily obtained by patient-clinician interaction; can direct clinical management decisions	May be influenced by patient-clinician interaction
VHI	Voice Handicap Index	Thirty-question patient-reported assessment of voice-related quality of life regarding impact of voice disorder on the patient's physical, functional, and emotional well-being	Validated assessment of patient perceived impact of voice deficit on quality of life; theoretically free from clinician influence. Easily obtained by simple questionnaire. Valuable outcomes measure	Not widely used or familiar to most surgeons

(continued)

Table VI. (*continued*)

Abbreviation	*Full term*	*Description*	*Benefits*	*Limitations*
NVO	Negative Voice Outcome	Novel comprehensive classification of voice outcomes after thyroidectomy that accounts for neurogenic and functionally relevant non-neurogenic voice outcomes by incorporating VCH, CAPE-V, DSI, and VHI in addition to VLS findings	Comprehensive diagnostic criteria that aims to capture patients with functionally relevant adverse voice after thyroidectomy despite frequent difficulties in determining precise anatomic cause	New, yet to be validated by other investigators. Requires comprehensive voice testing and expertise not universally available. Not inclusive of all acoustic measures

discrepancy may be that, in this study, loudness was not controlled; speakers were instructed to speak the sustained vowels and sentences in a comfortable, typical voice. This consideration is relevant because when adults phonate at a "comfortable level," they often produce sounds that occur at less than the typical intensity (eg, 75 dB) level.[28] A recent study by Brockerman et al[29] showed that the quieter a man's phonation, the greater his level of vocal jitter. As jitter levels increase, the DSI score decreases; thus, not controlling for loudness may have contributed to the effect observed according to patient sex in this study. Another contributory factor may be the relatively greater proportion of current and former smokers among males in this population.

Hakkesteegt et al[22] suggested that a change in DSI of 2.49 or more within an individual patient is clinically relevant. Although we have shown that change in DSI from preoperative values is predictive of long-term voice problems, we would advocate that the individual clinician should judge at which threshold change in DSI should trigger referrals for comprehensive voice evaluations. Sensitivity, specificity, and predictive values change expectedly across various DSI thresholds. For example, in our practice, we would seek a threshold with high sensitivity at the price of specificity and positive predictive value so as to help ensure that we do not miss patients who might ultimately benefit from comprehensive voice referrals and therapy. Admittedly, this approach will result in a relatively greater rate of referrals to voice specialists. In contrast, setting the DSI change threshold at –3.5 in our study would result in only 18% of our patients being referred for comprehensive voice testing and therapy, of whom 55% (positive predictive value) were identified ultimately as having an NVO.

We have previously identified that changes in the VHI and CAPE-V at the first postoperative visit relative to the pre-operative baseline are predictive of dysphonia in patients after thyroidectomy.[30] In that study, we defined dysphonia as VLS abnormalities in the presence of a patient-reported voice-specific symptom(s). In the current study, we could not compare the predictive value of the DSI relative to the VHI or CAPE-V because VHI and CAPE-V measurements were included in our revised, more inclusive diagnostic criteria for NVO. It should also be noted that the current CAPE-V data were collected in a randomized, blinded fashion, and the CAPE-V data previously reported were derived by 1 voice expert during the data-collection session. These values differ somewhat, as reported by Solomon et al.[25]

DSI measurements were inversely correlated with VHI and CAPE-V measurements. None of the statistically significant correlations, however, reached a level of correlation that we would deem to be compelling clinically. This finding suggests that the DSI calculation is not redundant with other assessments of voice quality; rather, the DSI may add additional functional voice data to a comprehensive assessment of voice. This concept is important to note, because DSI is purported to be a measure of dysphonia, and therefore it should correlate strongly with perceptual measures of dysphonia, such as the CAPE-V. Three of the 4 parameters of the DSI (quietest phonation, highest pitch, and MPT) are described more appropriately as measures of voice function rather than measures of dysphonia, which leads us to question its face validity or realism as a measure of vocal

quality. Our contention is that the DSI may serve as a complementary measure to both perceived dysphonia (CAPE-V) and functional, voice-related quality of life (VHI).

Most patients in this study with NVOs at the 6-month postoperative time point were classified as having subjective NVO, based on patient voice-specific complaints in the absence of structural or physiologic laryngeal impairment. Some of those classified as subjective NVO may have had undiagnosed injuries to the external branch of superior laryngeal nerve (EBSLN), the laryngoscopic findings of which are subtle typically.[31] Unfortunately, confirmatory cricothyroid electromyographic (EMG; considered to be the gold standard for diagnosis of EBSLN injury) was not performed in this study. Another explanation for the relatively high percentage of subjective versus objective NVO in this study is that VLS alone is insufficient in diagnosing all possible clinically relevant adverse vocal outcomes after thyroidectomy. Although considered to provide physiologic data, VLS is also subjective, because the results are based on visual-perceptual ratings, and optical illusions can occur.[32] Changes in DSI might help to identify these hard-to-diagnose patients, thus adding objective validation to voice-specific patient symptoms that might otherwise be regarded as "functional" when unsubstantiated by other means of objective diagnostic evaluation.

This study has limitations that deserve mention. The cohort is relatively small and few patients suffered a durable (6-month) NVO by our diagnostic criteria. The sample size of our study limits our ability to define the test's predictive values precisely at differing thresholds of change from baseline. As mentioned previously, laryngeal EMG was not a part of this study protocol. It is possible that some of our subjective NVO patients actually had unsubstantiated EBSLN injury. Last, voice recordings from which jitter was calculated in this study did not control for loudness. As mentioned previously, future studies intending to use jitter as a valid and reliable parameter in the DSI must control for loudness.

In conclusion, we have characterized changes in the DSI occurring in patients undergoing thyroidectomy. Changes in the DSI at the first postoperative visit relative to pre-operative baseline are highly predictive of patients who will experience long-term voice dysfunction after thyroidectomy. We suggest that these patients warrant comprehensive diagnostic and therapeutic interventions early in the postoperative period.

REFERENCES

1. Debruyne F, Ostyn F, Delaere P, Wellens W. Acoustic analysis of the speaking voice after thyroidectomy. J Voice 1997;11:479-82.
2. de Pedro Netto I, Fae A, Vartanian JG, Barros AP, Correia LM, Toledo RN, et al. Voice and vocal self-assessment after thyroidectomy. Head Neck 2006;28:1106-14.
3. Hong KH, Kim YK. Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. Otolaryngol Head Neck Surg 1997;117:399-404.
4. Lombardi CP, Raffaelli M, D'Alatri L, Marchese MR, Rigante M, Paludetti G, et al. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. Surgery 2006;140:1026-34.
5. McIvor NP, Flint DJ, Gillibrand J, Morton RP. Thyroid surgery and voice-related outcomes. Aust N Z J Surg 2000;70:179-83.
6. Sinagra DL, Montesinos MR, Tacchi VA, Moreno JC, Falco JE, Mezzadri NA, et al. Voice changes after thyroidectomy without recurrent laryngeal nerve injury. J Am Coll Surg 2004;199:556-60.
7. Soylu L, Ozbas S, Uslu HY, Kocak S. The evaluation of the causes of subjective voice disturbances after thyroid surgery. Am J Surg 2007;194:317-22.
8. Stojadinovic A, Shaha A, Orlikoff RF, Nissan A, Kornak MF, Singh B, et al. Prospective functional voice assessment in patients undergoing thyroid surgery. Ann Surg 2002;236:823-32.
9. Filho JG, Kowalski LP. Surgical complications after thyroid surgery performed in a cancer hospital. Otolaryngol Head Neck Surg 2005;132:490-4.
10. Roy AD, Gardiner RH, Niblock WM. Thyroidectomy and the recurrent laryngeal nerve. Lancet 1956;270:988-90.
11. Cernea C, Ferraz A, Fulani J, Monteiro S, Nishio S, Hojaij FC, et al. Identication of the external branch of the superior laryngeal nerve during thyroidectomy. Am J Surg 1992;164:634-8.
12. Hurtado-Lopez LM, Pacheco-Alvarez MI, Montes-Castillo Mde L, Zaldivar-Ramirez FR. Importance of the intraoperative identification of the external branch of the superior laryngeal nerve during thyroidectomy: electromyographic evaluation. Thyroid 2005;15:449-54.
13. Pereira JA, Girvent M, Sancho JJ, Parada C, Sitges-Serra A. Prevalence of long-term upper aero-digestive symptoms after uncomplicated bilateral thyroidectomy. Surgery 2003;133:318-22.
14. Kark AE, Kissin MW, Auerbach R, Meikle M. Voice changes after thyroidectomy: role of the external laryngeal nerve. Br Med J 1984;289:1412-5.
15. Henry LR, Solomon NP, Howard R, Gurevich-Uvena J, Horst L, Coppit G, et al. The functional impact on voice of sternothyroid muscle division during thyroidectomy. Ann Surg Onc 2008;15:2027-33.
16. Wuyts FL, De Bodt MS, Molenbergs G, Remacle M, Heylen L, Millet B, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. J Speech Lang Hearing Res 2000;43:796-809.
17. Hakkesteegt MM, Brocaar MP, Wieringa MH, Feenstra L. The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. J Voice 2008;22:138-45.
18. Van Lierde KM, De Ley S, Clement G, De Bodt M, Van Cauwenberge P. Outcome of laryngeal manual therapy in four Dutch adults with persistent moderate to severe vocal hyperfunction: a pilot study. J Voice 2004;18:467-74.

19. Timmermans B, De Bodt MS, Wuyts FL, Van de Heyning PH. Training outcomes in future professional voice users after 18 months of voice training. Folia Phoniatr Logop 2004;56:120-9.
20. Timmermans B, De Bodt M, Wuyts F, Van de Heyning P. Voice quality change in future professional voice users after 9 months of voice training. Eur Arch Othorhinolaryngol 2004;261:1-5.
21. Hakkesteegt MM, Brocaar MP, Wieringa MH, Feenstra L. Influence of age and gender on the dysphonia severity index. Folia Phoniatri Logop 2006;58:264-73.
22. Hakkesteegt MM, Wieringa MH, Brocaar MP, Mulder PGH, Feenstra L. The interobserver and test-retest variability of the Dysphonia Severity Index. Folia Phoniatr Logop 2008;60:86-90.
23. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. Am J Sp Lang Path 2009;18:124-32.
24. Jacobson B, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger M, et al. The Voice Handicap Index (VHI): development and validation. Am J Speech-Language Path 1997;6:66-9.
25. Solomon NP, Helou LB, Stojadinovic A. Clinical versus laboratory ratings of voice using the CAPE-V. J Voice. In press.
26. Awan SN, Alphonso VA. Effects of smoking on respiratory capacity and control. Clin Linguist Phon 2007;21:623-36.
27. Garnett JD. Tobacco and laryngeal pathology. W V Med J 2001;97:13-6.
28. Brown WS, Morris RJ, Murry T. Comfortable effort level revisited. J Voice 1996;10:299-305.
29. Brockerman M, Storck C, Carding PN, Drinnan MJ. Voice loudness and gender effects on jitter and shimmer in healthy adults. J Speech Lang Hear Res 2008;51:1152-60.
30. Stojadinovic A, Henry LR, Solomon NP, Gurevich-Uvena J, Makashay MJ, Coppit GL, et al. Prospective trial of voice outcomes following thyroidectomy: Evaluation of patient-reported and clinician-determined voice assessments in identifying post-thyroidectomy dysphonia. Surgery 2008;143:732-42.
31. Sataloff RT, Praneetvatakul P, Heuer RJ, Hawkshaw MJ, Heman-Ackah YD, Schneider SM, et al. Laryngeal electromyography: clinical application. J Voice. Epub ahead of print.
32. Krantz JH. Did I really see that? The complex relationship between the visual stimulus and visual perception. J Voice 2008;22:520-32.